

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
21 November 2002 (21.11.2002)

PCT

(10) International Publication Number
WO 02/093344 A1

(51) International Patent Classification⁷: **G06F 3/00**

(21) International Application Number: **PCT/IB02/01666**

(22) International Filing Date: **14 May 2002 (14.05.2002)**

(25) Filing Language: **English**

(26) Publication Language: **English**

(30) Priority Data:
01201799.2 14 May 2001 (14.05.2001) EP

(71) Applicant (for all designated States except US): **KONINKLIJKE PHILIPS ELECTRONICS N.V. [NL/NL];**
Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **STIENSTRA, Mar-
celle, A. [NL/NL];** Prof. Holstlaan 6, NL-5656 AA Eind-
hoven (NL).

(74) Agent: **GROENENDAAL, Antonius, W., M.;** Interna-
tionaal Octrooibureau B.V., Prof. Holstlaan 6, NL-5656
AA Eindhoven (NL).

(81) Designated States (*national*): **CN, JP, KR, US.**

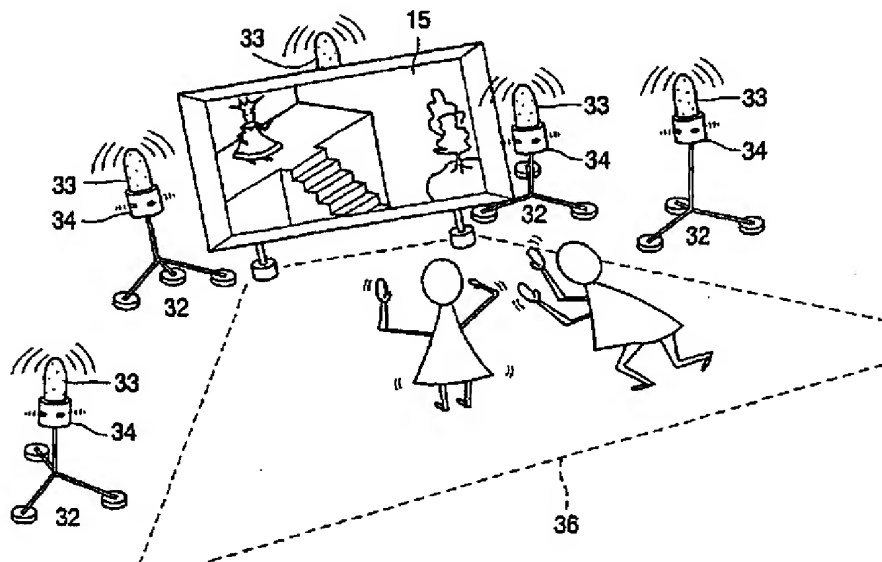
(84) Designated States (*regional*): **European patent (AT, BE,
CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC,
NL, PT, SE, TR).**

Published:

- with international search report
- before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: **DEVICE FOR INTERACTING WITH REAL-TIME STREAMS OF CONTENT**



(57) Abstract: An end-user system (10) for transforming real-time streams of content into an output presentation includes a user interface (30) that allows a user to interact with the streams. The user interface (30) includes sensors (32a-f) that monitor an interaction area (36) to detect movements and/or sounds made by a user. The sensors (32a-f) are distributed among the interaction area (36) such that the user interface (30) can determine a three-dimensional location within the interaction area (36) where the detected movement or sound occurred. Different streams of content can be activated in a presentation based on the type of movement or sound detected, as well as the determined location. The present invention allows a user to interact with and adapt the output presentation according to his/her own preferences, instead of merely being a spectator.

WO 02/093344 A1

Device for interacting with real-time streams of content

The present invention relates to a system and method for receiving and displaying real-time streams of content. Specifically, the present invention enables a user to interact with and personalize the displayed real-time streams of content.

5

Storytelling and other forms of narration have always been a popular form of entertainment and education. Among the earliest forms of these are oral narration, song, written communication, theater, and printed publications. As a result of the technological advancements of the nineteenth and twentieth century, stories can now be broadcast to large numbers of people at different locations. Broadcast media, such as radio and television, allow storytellers to express their ideas to audiences by transmitting a stream of content, or data, simultaneously to end-user devices that transforms the streams for audio and/or visual output.

Such broadcast media are limited in that they transmit a single stream of content to the end-user devices, and therefore convey a story that cannot deviate from its predetermined sequence. The users of these devices are merely spectators and are unable to have an effect on the outcome of the story. The only interaction that a user can have with the real-time streams of content broadcast over television or radio is switching between streams of content, i.e., by changing the channel. It would be advantageous to provide users with more interaction with the storytelling process, allowing them to be creative and help determine how the plot unfolds according to their preferences, and therefore make the experience more enjoyable.

At the present time, computers provide a medium for users to interact with real-time streams of content. Computer games, for example, have been created that allow users to control the actions of a character situated in a virtual environment, such as a cave or a castle. A player must control his/her character to interact with other characters, negotiate obstacles, and choose a path to take within the virtual environment. In on-line computer games, streams of real-time content are broadcast from a server to multiple personal computers over a network, such that multiple players can interact with the same characters,

obstacles, and environment. While such computer games give users some freedom to determine how the story unfolds (i.e., what happens to the character), the story tends to be very repetitive and lacking dramatic value, since the character is required to repeat the same actions (e.g. shooting a gun), resulting in the same effects, for the majority of the game's duration.

Various types of children's educational software have also been developed that allows children to interact with a storytelling environment on a computer. For example, LivingBooks® has developed a type of "interactive book" that divides a story into several scenes, and after playing a short animated clip for each scene, allows a child to manipulate various elements in the scene (e.g., "point-and-click" with a mouse) to play short animations or gags. Other types of software provide children with tools to express their own feelings and emotions by creating their own stories. In addition to having entertainment value, interactive storytelling has proven to be a powerful tool for developing the language, social, and cognitive skills of young children.

However, one problem associated with such software is that children are usually required to using either a keyboard or a mouse in order to interact. Such input devices must be held in a particular way and require a certain amount of hand-eye coordination, and therefore may be very difficult for younger children to use. Furthermore, a very important part of the early cognitive development of children is dealing with their physical environment. An interface that encourages children to interact by "playing" is advantageous over the conventional keyboard and mouse interface, because it is more beneficial from an educational perspective, it is more intuitive and easy to use, and playing provides a greater motivation for children to participate in the learning process. Also, an interface that expands the play area (i.e., area in which children can interact), as well as allowing children to interact with objects they normally play with, can encourage more playful interaction.

ActiMates™ Barney™ is an interactive learning product created by Microsoft Corp.®, which consists of a small computer embedded in an animated plush doll. A more detailed description of this product is provided in the paper, E. Strommen, "When the Interface is a Talking Dinosaur: Learning Across Media with ActiMates Barney," Proceedings of CHI '98, pages 288-295. Children interact with the toy by squeezing the doll's hand to play games, squeezing the doll's toe to hear songs, and covering the doll's eyes to play "peek-a-boo." ActiMates Barney can also receive radio signals from a personal computer and coach children while they play educational games offered by ActiMates

software. While this particular product fosters interaction among children, the interaction involves nothing more than following instructions. The doll does not teach creativity or collaboration, which are very important in the developmental learning, because it does not allow the child to control any of the action.

5 CARESS (Creating Aesthetically Resonant Environments in Sound) is a project for designing tools that motivate children to develop creativity and communication skills by utilizing a computer interface that converts physical gestures into sound. The interface includes wearable sensors that detect muscular activity and are sensitive enough to detect intended movements. These sensors are particularly useful in allowing physically
10 challenged children to express themselves and communicate with others, thereby motivating them to participate in the learning process. However, the CARESS project does not contemplate an interface that allows the user any type of interaction with streams of content.

15 It is an object of the present invention to allow users to interact with real-time streams of content received at an end-user device. This object is achieved according to the invention in a user interface as claimed in claim 1. Real-time streams of content are transformed into a presentation that is output to the user by an output device, such as a television or computer display. The presentation conveys a narrative whose plot unfolds
20 according to the transformed real-time streams of content, and the user's interaction with these streams of content help determine the outcome of the story by activating or deactivating streams of content, or by modifying the information transported in these streams. The user interface allows users to interact with the real-time streams of content in a simple, direct, and intuitive manner. The interface provides users with physical, as well as mental, stimulation
25 while interacting with real-time streams of content.

One embodiment of the present invention is directed to a system that transforms real-time streams of content into a presentation to be output and a user interface through which a user activates or deactivates streams of content within the presentation.

30 In another embodiment of the present invention, the user interface includes at least one motion detector that detects movements or gestures made by a user. In this embodiment, the detected movements determine which streams of content are activated or deactivated.

In another embodiment, the user interface includes a plurality of motion sensors that are positioned in such a way as to detect and differentiate between movements made by one or more users at different locations within a three-dimensional space.

5 In another embodiment of the present invention, a specific movement or combination of specific movements are correlated to a specific stream of content. When the motion sensors of the user interface detect a specific movement or combination of movements made by the user, the corresponding stream of content is either activated or deactivated.

10 In another embodiment of the present invention, the user interface includes a plurality of sensors that detect sounds. In this embodiment, the detected sounds determine which streams of content are activated or deactivated.

15 In another embodiment of the present invention, the user interface includes a plurality of sound-detecting sensors that are positioned in such a way as to detect and differentiate between specific sounds made by one or more users at different locations within a three-dimensional space.

20 In another embodiment the user interface includes a combination of motion sensors and sound-detecting sensors. In this embodiment, streams of content are activated according to a detected movement or sound made by a user, or a combination of detected movements and sounds.

These and other embodiments of the present invention will become apparent from and elucidated with reference to the following detailed description considered in connection with the accompanying drawings.

25 It is to be understood that these drawings are designed for purposes of illustration only and not as a definition of the limits of the invention for which reference should be made to the appending claims.

Fig. 1 is a block diagram illustrating the configuration of a system for transforming real-time streams of content into a presentation.

30 Fig. 2 illustrates the user interface of the present invention according to an exemplary embodiment.

Figs. 3A and 3B illustrate a top view and a side view, respectively, of the user interface.

Fig. 4 is a flowchart illustrating the method whereby real-time streams of content can be transformed into a narrative.

5 Referring to the drawings, Fig. 1 shows a configuration of a system for transforming real-time streams of content into a presentation, according to an exemplary embodiment of the present invention. An end-user device 10 receives real-time streams of data, or content, and transforms the streams into a form that is suitable for output to a user on output device 15. The end-user device 10 can be configured as either hardware, software
10 being executed on a microprocessor, or a combination of the two. One possible implementation of the end-user device 10 and output device 15 of the present invention is as a set-top box that decodes streams of data to be sent to a television set. The end-user device 10 can also be implemented in a personal computer system for decoding and processing data streams to be output on the CRT display and speakers of the computer. Many different
15 configurations are possible, as is known to those of ordinary skill in the art.

The real-time streams of content can be data streams encoded according to a standard suitable for compressing and transmitting multimedia data, for example, one of the Moving Picture Experts Group (MPEG) series of standards. However, the real-time streams of content are not limited to any particular data format or encoding scheme. As shown in
20 Fig. 1, the real-time streams of content can be transmitted to the end-user device over a wire or wireless network, from one of several different external sources, such as a television broadcast station 50 or a computer network server. Alternatively, the real-time streams of data can be retrieved from a data storage device 70, e.g. a CD-ROM, floppy-disc, or Digital Versatile Disc (DVD), which is connected to the end-user device.

25 As discussed above, the real-time streams of content are transformed into a presentation to be communicated to the user via output device 15. In an exemplary embodiment of the present invention, the presentation conveys a story, or narrative, to the user. Unlike prior art systems that merely convey a story whose plot is predetermined by the real-time streams of content, the present invention includes a user interface 30 that allows the
30 user to interact with a narrative presentation and help determine its outcome, by activating or deactivating streams of content associated with the presentation. For example, each stream of content may cause the narrative to follow a particular storyline, and the user determines how the plot unfolds by activating a particular stream, or storyline. Therefore, the present invention allows the user to exert creativity and personalize the narrative according to his/her

own wishes. However, the present invention is not limited to transforming real-time streams of content into a narrative to be presented to the user. According to other exemplary embodiments of the present invention, the real-time streams can be used to convey songs, poems, musical compositions, games, virtual environments, adaptable images, or any other
5 type of content with which the user can adapt according to his/her personal wishes.

As mentioned above, Fig. 2 shows in detail the user interface 30 according to an exemplary embodiment, which includes a plurality of sensors 32 distributed among a three-dimensional area in which a user interacts. The interaction area 36 is usually in close proximity to the output device 15. In an exemplary embodiment, each sensor 32 includes
10 either a motion sensor 34 for detecting user movements or gestures, a sound-detecting sensor 33 (e.g., a microphone) for detecting sounds made by a user, or a combination of both a motion sensor 34 and a sound-detecting sensor 33 (Fig. 2 illustrates sensors 32 that include such a combination).

The motion sensor 34 may comprise an active sensor that injects energy into
15 the environment to detect a change caused by motion. One example of an active motion sensor comprises a light beam that is sensed by a photosensor. The photosensor is capable of detecting a person or object moving across, and thereby interrupting, the light beam by detecting a change in the amount of light being sensed. Another type of active motion sensor uses a form of radar. This type of sensor sends out a burst of microwave energy and waits for
20 the reflected energy to bounce back. When a person comes into the region of the microwave energy, the sensor detects a change in the amount of reflected energy or in the time it takes for the reflection to arrive. Other active motion sensors similarly use reflected ultrasonic sound waves to detect motion.

Alternatively, the motion sensor 34 may comprise a passive sensor, which
25 detects infrared energy being radiated from a user. Such devices are known as PIR detectors (Passive InfraRed) and are designed to detect infrared energy having a wavelength between 9 and 10 micrometers. This range of wavelength corresponds to the infrared energy radiated by humans. Movement is detected according to a change in the infrared energy being sensed, caused by a person entering or exiting the field of detection. PIR sensors typically have a
30 very wide angle of detection (up to, and exceeding, 175 degrees).

Of course, other types of motion sensors may be used in the user interface 30, including wearable motion sensors and video motion detectors. Wearable motion sensors may include virtual reality gloves, sensors that detect electrical activity in muscles, and sensors that detect the movement of body joints. Video motion detectors detect movement in

images taken by a video camera. One type of video motion detector detects sudden changes in the light level of a selected area of the images to detect movement. More sophisticated video motion detectors utilize a computer running image analysis software. Such software may be capable of differentiating between different facial expressions or hand gestures made
5 by a user.

The user interface 30 may incorporate one or more of the motion sensors described above, as well as any other type of sensor that detects movement that is known in the art.

The sound-detecting sensor 33 may include any type of transducer for
10 converting sound waves into an electrical signal (such as a microphone). The electrical signals picked up by the sound sensors can be compared to a threshold signal to differentiate between sounds made by a user and environmental noise. Further, the signals may be amplified and processed by an analog device or by software executed on a computer to detect sounds having particular frequency pattern. Therefore, the sound-detecting sensor 34 may
15 differentiate between different types of sounds, such as stomping feet and clapping hands.

The sound-detecting sensor 33 may include a speech recognition system for recognizing certain words spoken by a user. The sound waves may be converted into amplified electrical signals that are processed by an analog speech recognition system, which is capable of recognizing a limited vocabulary of words; else, the converted electrical signals
20 may be digitized and processed by speech recognition software, which is capable of recognizing a larger vocabulary of words.

The sound-detecting sensor 33 may comprise one of a variety of embodiments and modifications, as is well known to those skilled in the art. According to an exemplary embodiment, the user interface 30 may incorporate one or more sound-detecting sensors 34
25 taking on one or more different embodiments.

Figs. 3A and 3B illustrate an exemplary embodiment of the user interface 30, in which a plurality of sensors 32a-f that are positioned around an interactive area 36, in which a user interacts. The sensors 32a-f are positioned so that the user interface 30 not only detects whether or not a movement or sound has been made by the user within interaction
30 area 36, but also determines a specific location in interaction area 36 that the movement or sound was made. As shown in Figs. 3A and 3B, the interaction area 36 can be divided into a plurality of areas in three-dimensions. Specifically, Fig. 3A illustrates an overhead view of the user interface 30, where the two-dimensional plane of the interaction area 36 is divided into quadrants 36a-d. Fig. 3B illustrates a side view of the user interface 30, where the

interaction area is further divided according to a third dimension (vertical) into areas 36U and 36L. In the embodiment shown in Figs. 3A and 3B, the interaction area 36 can be divided into eight three-dimensional areas: (36a, 36U), (36a, 36L), (36b, 36U), (36b, 36L), (36c, 36U), (36c, 36L), (36d, 36U), and (36d, 36L).

5 According to this embodiment, the user-interface 30 is able to determine a three-dimensional location in which a movement or sound is detected, because multiple sensors 32a-f are positioned around the interaction area 36. Figure 32A shows that sensors 32a-f are positioned such that a movement or sound made in quadrants 36a or 36c will produce a stronger detection signal in sensors 32a, 32b, and 32f than in sensors 32c, 32d, and
10 32e. Likewise, a sound or movement made in quadrants 36c or 36d will produce a stronger detection signal in sensors 32f and 32e than in sensors 32b and 32c.

Figure 3B also shows that sensors 32a-f have located at various elevations. For example, sensors 32b, 32f, and 32d will more strongly detect a movement or noise made close to the ground than will sensors 32a, 32c, and 32e.

15 The user interface 30 can therefore determine in which three-dimensional area the movement or sound was made based on the position of each sensor, as well as the strength the signal generated by the sensor. As an example, an embodiment in which sensors 32a-f each contain a PIR sensor will be described below in connection with Figs. 3A and 3B.

When a user waves his hand in location (36b, 36U), each PIR sensor 34 of
20 sensors 32a-f may detect some amount change in the infrared energy sensed. However, the PIR sensor of sensor 32c will sense the greatest amount of change because of its proximity to the movement. Therefore, sensor 32c will output the strongest detection signal, and the user-interface can determine the three-dimensional location in which the movement was made, by determining which three-dimensional location is closest to sensor 32c.

25 Similarly, the location of sounds made by users in the interaction area 36 can be determined according to the respective locations and magnitude of detection signals output by the sound-detecting sensors 33 in sensors 32a-f.

Figs. 3A and 3B shows an exemplary embodiment and should not be construed as limiting the present invention. According to another exemplary embodiment,
30 the user-interface 30 may include a video motion detector that includes image-processing software for analyzing the video image to determine the type and location of movement within an interaction area 36. In another exemplary embodiment, the user interface may also comprise a grid of piezoelectric cables covering the floor of the interaction area 36 that senses the location and force of footsteps made by a user.

In an exemplary embodiment, the end-user device 10 determines which streams of content should be activated or deactivated in the presentation, based on the type of movements and/or sounds detected by the user interface 30. In this embodiment, each stream of content received by the end-user device may include control data that links the stream to a particular gesture or movement. For example, the stomping of feet may be linked to a stream of content that causes a character in the narrative to start walking or running. Similarly, a gesture that imitates the use of a device or tool (e.g. a scooping motion for using a shovel) may be linked to a stream that causes the character to use that device or tool.

In a further exemplary embodiment, a user can imitate a motion or a sound being output in connection with a particular activated stream of content, in order to deactivate the stream. Conversely, the user can imitate a motion or sound of a particular stream of content to select that stream for further manipulation by the user.

In another exemplary embodiment, a particular stream of content may be activated according to a specific word spoken or a specific type of sound made by one or more users. Similar to the previously described embodiment, each received stream of content may include control data for linking it to a specific word or sound. For example, by speaking the word of an action (e.g., "run"), a user may cause the character of a narrative to perform the corresponding action. By making a sound normally associated with an object, a user may cause that object to appear on a screen or to be used by a character. For example, by saying "pig" or "oink," the user may cause a pig to appear.

In another exemplary embodiment, the stream of content may include control data that links the stream to a particular location in which a movement or sound is made. For example, if a user wants a character to move in a particular direction, the user can point to the particular direction. The user interface 30 will determine the location that the user moved his/her hand to, and send the location information to the end-user device 10, which activates the stream of content that causes the character to move in the corresponding direction.

In another exemplary embodiment, the stream of content may include control data to link the stream to a particular movement or sound, and the end-user device 10 may cause the stream to be displayed at an on-screen location corresponding to the location where the user makes the movement or sound. For example, when a user practices dance steps, each step taken by the user may cause a footprint to be displayed on a screen location corresponding to the location of the actual step within the interaction area.

According to another exemplary embodiment, the user interface 30 determines not only the type of movement or sound made by the user, but also the manner in which the

movement or sound was made. For example, the user interface can determine how loudly a user issues an oral command by analyzing the magnitude of the detected sound waves. Also, the user interface 30 may determine the amount of force or speed with which a user makes a gesture. For example, active motion sensors that measure reflected energy (e.g., radar) can
5 detect the speed of movement. In addition, pressure based sensors, such as a grid of piezoelectric cables, can be used to detect the force of certain movements.

In the above embodiment, the manner in which a stream of content is output depends on the manner in which a user makes the movement or sound that activates the stream. For example, the loudness of a user's singing can be used to determine how long a
10 stream remains visible on screen. Likewise, the force with which the user stomps his feet can be used to determine how rapidly a stream moves across the screen.

Another exemplary embodiment of the present invention, a stream of content is activated or deactivated according to a series or combination of movements and/or sounds. This embodiment can be implemented by including control data in a received stream that
15 links the stream to a group of movements and/or sounds. Possible implementations of this embodiment include activating or deactivating a stream when the sensors 32 detect a set of movements and/or sound in a specific sequence or within a certain time duration.

According to another exemplary embodiment, control data may be provided with the real-time streams of content received at the end-user device 10 that automatically
20 activates or deactivates certain streams of content. This allows the creator(s) of the real-time streams to have some control over what streams of content are activated and deactivated. In this embodiment, the author(s) of a narrative has a certain amount of control as to how the plot unfolds by activating or deactivating certain streams of content according to control data within the transmitted real-time streams of content.

25 In another exemplary embodiment of the present invention, when multiple users are interacting with the present invention at the same time, the user-interface 30 can differentiate between sounds or movements made by each user. Therefore, each user may be given the authority to activate or deactivate different streams of content by the end-user device. Sound-detecting sensors 33 may be equipped with voice recognition hardware or
30 software that allows the user-interface to determine which user speaks a certain command. The user interface 30 may differentiate between movements of different users by assigning a particular section of the interaction area 36 to each user. Whenever a movement is detected at a certain location of the interaction area 36, the user interface will attribute the movement

to the assigned user. Further, video motion detectors may include image analysis software that is capable of identifying a user that makes a particular movement.

In the above embodiment, each user may control a different character in an interactive narrative presentation. Control data within a stream of content may link the stream to the particular user to who may activate or deactivate it. Therefore, only the user who controls a particular character can activate or deactivate streams of content relating to that character.

In another exemplary embodiment, two or more streams of content activated by two or more different users may be combined into a single stream of content. For example, after each user activates a stream of content, they can combine the activated streams by issuing an oral command (e.g., "combine") or by making a particular movement (e.g., moving toward each other).

According to another exemplary embodiment, the user interface 30 may include one or more objects for user(s) to manipulate in order to activate or deactivate a stream. In this embodiment, a user causes the object to move and/or to make a particular sound, and the sensors 32 detect this movement and/or sound. For instance, the user will be allowed to kick or throw a ball, and the user interface 30 will determine the distance, direction, and/or velocity at which the ball traveled. Alternatively, the user may play a musical instrument, and the user interface will be able to detect the notes played by the user. Such an embodiment can be used to activate streams of content in a sports simulation game or in a program that teaches a user how to play a musical instrument.

As described above, an exemplary embodiment of the present invention is directed to an end-user device that transforms real-time streams of content into a narrative that is presented to the user through output device 15. One possible implementation of this embodiment is an interactive television system. The end-user device 10 can be implemented as a set-top box, and the output device 15 is the television set. The process by which a user interacts with such a system is described below in connection with the flowchart 100 of Fig. 4.

In step 110, the end-user device 10 receives a stream of data corresponding to a new scene of a narrative and immediately processes the stream of data to extract scene data. Each narrative presentation includes a series of scenes. Each scene comprises a setting in which some type of action takes place. Further, each scene has multiple streams of content associated therewith, where each stream of content introduces an element that affects the plot.

For example, activation of a stream of content may cause a character to perform a certain action (e.g., a prince starts walking in a certain direction), cause an event to occur that affects the setting (e.g., thunderstorm, earthquake), or introduce a new character to the narrative (e.g., frog). Conversely, deactivation of a stream of content may cause a character to stop performing a certain action (e.g., prince stops walking), terminate an event (e.g., thunderstorm or earthquake ends), or cause a character to depart from the story (e.g. frog hops away).

The activation or deactivation of a stream of content may also change an internal property or characteristic of an object in the presentation. For example, activation of a particular stream may cause the mood of a character, such as the prince, to change from happy to sad. Such a change may become evident immediately in the presentation (e.g., the prince's smile becomes a frown), or may not be apparent until later in the presentation. Such internal changes are not limited to characters, and may apply to any object that is part of the presentation, which contains some characteristic or parameter that can be changed.

In step 120, the set-top box decodes the extracted scene data. The setting is displayed on a television screen, along with some indication to the user that he/she must determine how the story proceeds by interacting with user interface 30. As a result, the user makes a particular movement or sound in the interaction area 36, as shown in step 130.

In step 140, the sensors 32 detect the movement(s) or sound(s) made by the user, and make a determination as to the type of movement or sound made. This step may include determining which user made the sound or movement, when multiple users are in the interaction area 36. In step 150, the set-top box determines which streams of content are linked to the determined movement or sound. This step may include examining the control data of each stream of content to determine whether the detected movement or sound is linked to the stream.

In step 160, the new storyline is played out on the television according to the activated/deactivated streams of content. In this particular example, each stream of content is an MPEG file, which is played on the television while activated.

In step 170, the set-top box determines whether the activated streams of content necessarily cause the storyline to progress to a new scene. If so, the process returns to step 110 to receive the streams of content corresponding to the new scene. However, if a new scene is not necessitated by the storyline, the set-top box determines whether the narrative has reached a suitable ending point in step 180. If this is not the case, the user is instructed to use the user interface 30 in order to activate or deactivate streams of content and

thereby continue the narrative. The flowchart of Fig. 4 and the corresponding description above is meant to describe an exemplary embodiment, and is in no way limiting.

The present invention provides a system that has many uses in the developmental education of children. The present invention promotes creativity and development of communication skills by allowing children to express themselves by
5 interacting with and adapting a presentation, such as a story. The present invention does not include a user interface that may be difficult to use for younger children, such as a keyboard and mouse. Instead, the present invention utilizes a user interface 30 that allows for basic, familiar sounds and movements to be linked to specific streams of contents. Therefore, the
10 child's interaction with the user interface 30 can be very "playful," providing children with more incentive to interact. Furthermore, streams of content can be linked with movements or sounds having a logical connection to the stream, thereby making interaction much more intuitive for children.

It should be noted, however, that the input device 30 of the present invention
15 is in no way limited in its use to children, nor is it limited to educational applications. The present invention provides an intuitive and stimulating interface to interact with many different kinds of presentations geared to users of all ages.

A user can have a variety of different types of interactions with the presentation by utilizing the present invention. As mentioned above, the user may affect the
20 outcome of a story by causing characters to perform certain types actions or by initiating certain events that affect the setting and all of the characters therein, such as a natural disaster or a weather storm. The user interface 30 can also be used to merely change details within the setting, such as changing the color of a building or the number of trees in a forest. However, the user is not limited to interacting with presentations that are narrative by nature.
25 The user interface 30 can be used to choose elements to be displayed in a picture, to determine the lyrics to be used in a song or poem, to play a game, to interact with a computer simulation, or to perform any type of interaction that permits self-expression of a user within a presentation. Furthermore, the presentation may comprise a tutoring program for learning physical skills (e.g., learn how to dance or swing a golf club) or verbal skills (e.g., learn how
30 to speak a foreign language or how to sing), in which the user can practice these skills and receive feedback from the program.

In addition, the user interface 30 of the present invention is not limited to an embodiment comprising motion and sound-detecting sensors 32 that surround and detect movements within a specified area. The present invention covers any type of user interface

in which the sensed movements of a user or object causes the activation or deactivation of streams of content. For example, the user interface 30 may include an object that contains sensors, which detect any type of movement or user manipulation of the object. The sensor signal may be transmitted from the object by wire or radio signals to the end-user device 10, which activates or deactivates streams of content as a result.

Furthermore, the present invention is not limited to detecting movements or sound made by a user in a specified interaction area 30. The present invention may comprise a sensor, such as a Global Positioning System (GPS) receiver, that tracks its own movement. In this embodiment, the present invention may comprise a portable end-user device 10 that activates received streams of content in order to display real-time data, such as traffic news, weather report, etc., corresponding to its current location.

The present invention has been described with reference to the exemplary embodiments. As will be evident to those skilled in the art, various modifications of this invention can be made or followed in light of the foregoing disclosure without departing from the scope of the claims.

CLAIMS:

1. A user interface (30) for interacting with a device that receives and transforms streams of content into a presentation to be output, comprising:
an interaction area (36);
at least one sensor (32) for detecting a movement or sound made by a user
5 within said interaction area (36),
wherein one or more streams of content are manipulated based on said detected movement or sound, and
wherein the presentation is controlled based on said manipulated streams of content.
- 10 2. The user interface (30) according to claim 1, wherein said at least one sensor (32) detects a movement made by the user,
and wherein a type of movement or sound corresponding to said detected movement or sound is determined by analyzing a detection signal from said at least one
15 sensor (32).
3. The user interface (30) according to claim 2, wherein a received stream of content is activated or deactivated in the presentation based on the determined type of movement or sound.
- 20 4. The user interface (30) according to claim 1, wherein said at least one sensor (32) includes a plurality of sensors, and
wherein detection signals from said plurality of sensors are analyzed to determine a location within said interaction area (36) in which said detected movement or
25 sound occurs.
5. The user interface (30) according to claim 4, wherein a received stream of content is activated or deactivated in the presentation based on said determined location.

6. The user interface (30) according to claim 1, wherein said at least one sensor (32) includes a sound-detecting sensor (33) connected to a speech recognition system, and wherein a received stream of content is detected based on a particular word being recognized by said speech recognition system.
- 5
7. The user interface (30) according to claim 1, wherein each of said at least one sensor (32) includes a motion sensor (34) and a sound-detecting sensor (33).
8. The user interface (30) according to claim 1, wherein said presentation
10 includes a narrative.
9. A process in a system for transforming streams of content into a presentation to be output, comprising:
- 15 detecting a movement or sound occurring within an interaction area (36);
manipulating one or more streams of content based on said detected movement or sound;
controlling said presentation based on the manipulated streams of content.
10. A system comprising:
- 20 an end-user device (10) for receiving and transforming streams of content into a presentation;
a user interface (30) including sensors (32) for detecting a movement or sound made by a user within an interaction area (36);
an output device (15) for outputting said presentation,
25 wherein said end-user device (10) manipulates said transformed streams of content based on said detected movement or sound, thereby controlling said presentation.

1/3

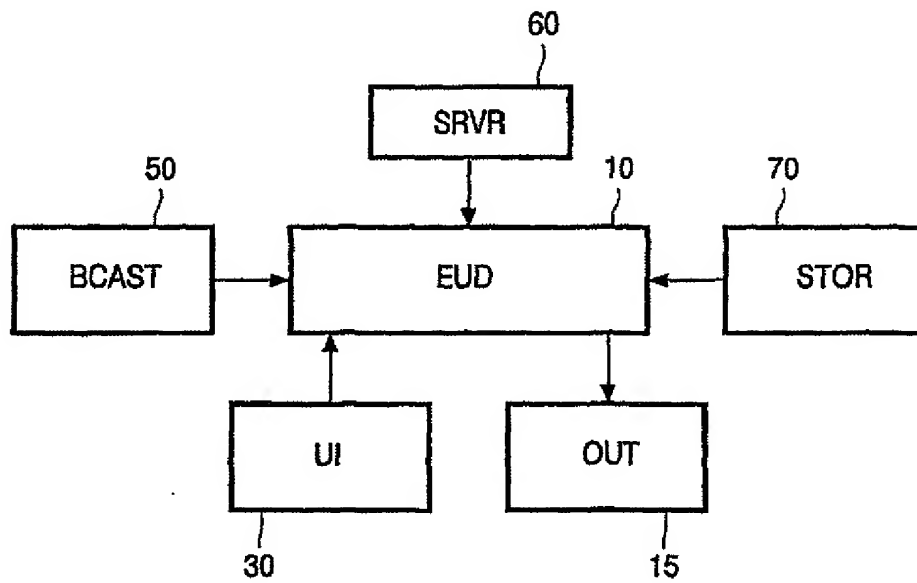


FIG. 1

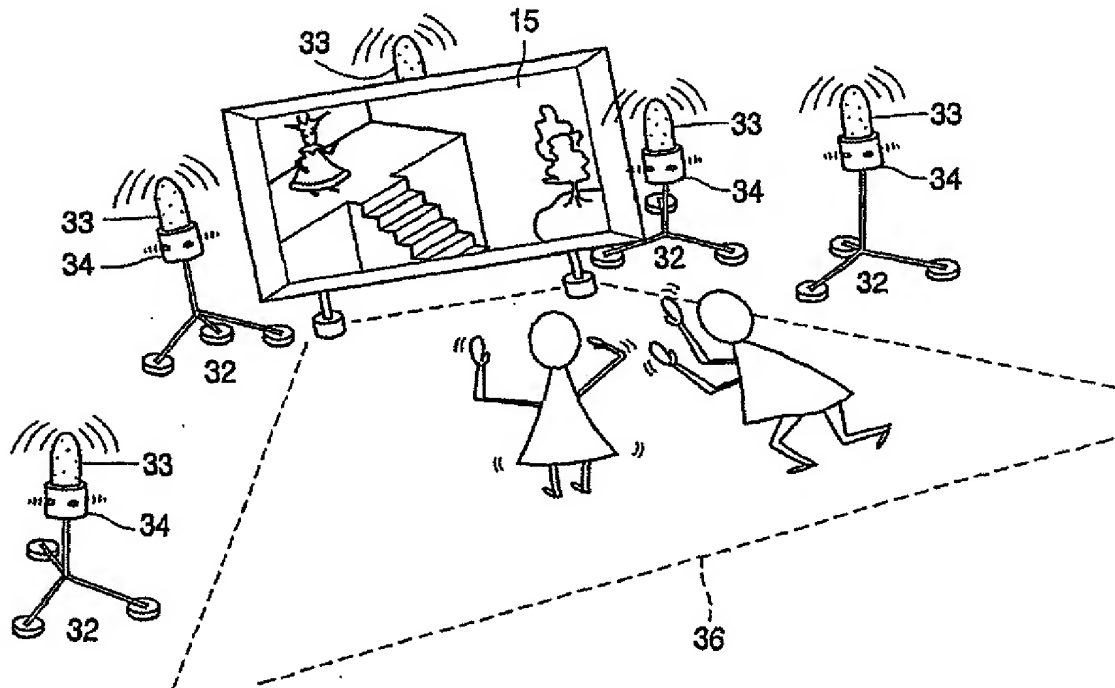


FIG. 2

2/3

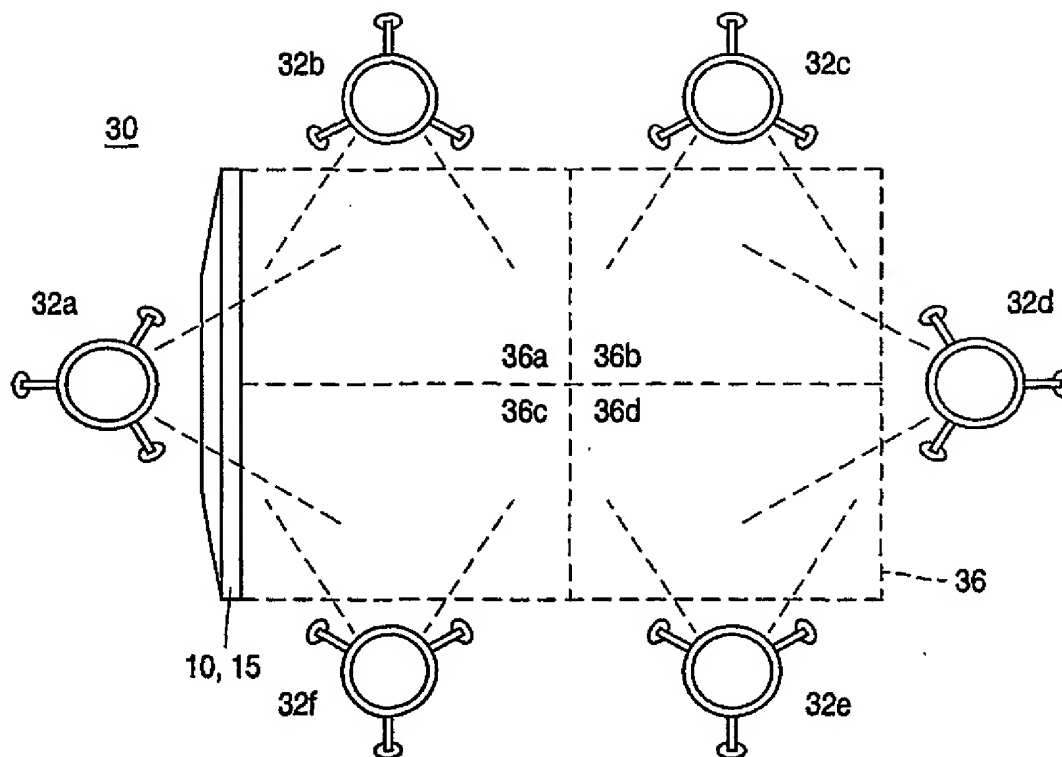


FIG. 3A

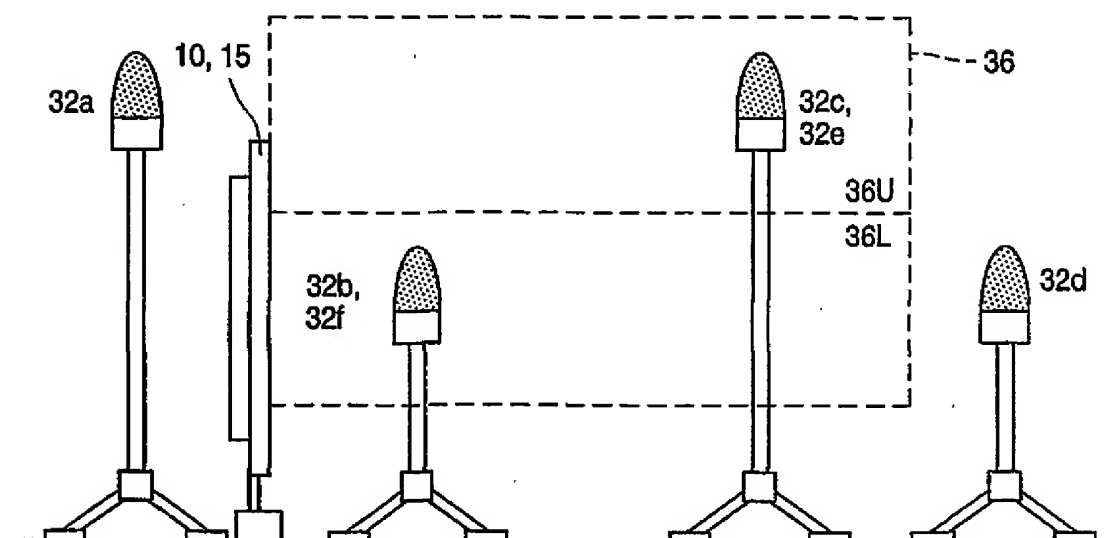


FIG. 3B

3/3

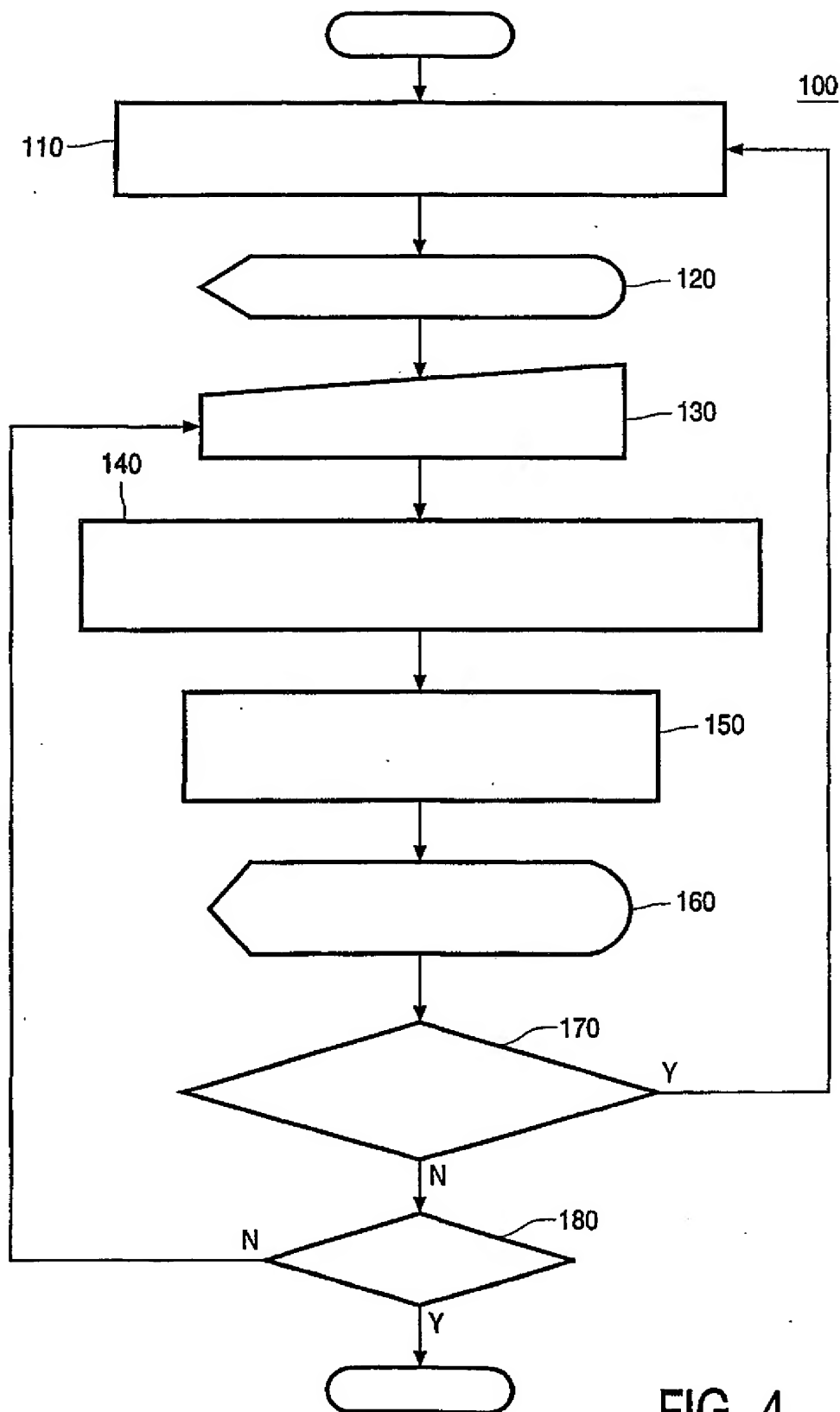


FIG. 4

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/IB 02/01666

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G06F3/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G06F 610H H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	US 4 569 026 A (BEST ROBERT M) 4 February 1986 (1986-02-04) abstract figure 13	1-3, 6, 8, 9 10
A	US 5 442 168 A (GURNER ET AL) 15 August 1995 (1995-08-15) abstract	1, 2, 4, 8-10
A	EP 0 654 779 A (PODESTA ET AL) 24 May 1995 (1995-05-24) abstract column 3, line 21 - line 30	1, 2, 4, 6, 7, 9, 10



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents:

A document defining the general state of the art which is not considered to be of particular relevance

E earlier document but published on or after the international filing date

L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

O document referring to an oral disclosure, use, exhibition or other means

P document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

Z document member of the same patent family

Date of the actual completion of the international search

4 October 2002

Date of mailing of the international search report

11/10/2002

Name and mailing address of the ISA

European Patent Office, P.B. 5010 Patentlaan 2
NL - 2280 HV Rijswijk
Tel: (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Dockhorn, H

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/IB 02/01666

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
US 4569026	A	04-02-1986	EP 0016314 A1	01-10-1980
			JP 1042702 B	14-09-1989
			JP 1556591 C	23-04-1990
			JP 55112000 A	29-08-1980
			US 4305131 A	08-12-1981
			US 4333152 A	01-06-1982
			US 4445187 A	24-04-1984
US 5442168	A	15-08-1995	IL 101016 A	14-05-1996
			BR 9300618 A	24-08-1993
			CA 2089932 A1	20-08-1993
			EP 0557111 A1	25-08-1993
			JP 2796028 B2	10-09-1998
			JP 6054935 A	01-03-1994
			US 5414256 A	09-05-1995
			US 5459312 A	17-10-1995
			US 5475214 A	12-12-1995
			US 5369270 A	29-11-1994
EP 0654779	A	24-05-1995	IT GE930098 A1	24-05-1995
			EP 0654779 A2	24-05-1995